

REMARKS

In the Office Action dated December 5, 2006, the Examiner rejects claims 1 through 13, 16 through 25, 27 through 36 and 38 through 46 as anticipated under 35 U.S.C. 102(e) by U.S. Patent No. 6,684,205 to Modha ("Modha"). The Examiner also rejects claims 13 and 47 as obvious under 35 U.S.C. 103(a) over Modha in view of the article entitled "Analysis of a Very Large AltaVista Query Log" by Silverstein, as well as claim 15 as obvious over Modha in view of the article entitled "Web Usage Mining: Discovery and Applications of Usage Patterns from Web Data" by Srivastava. Finally, the Examiner indicates that claim 26 stands objected to, but would be allowable if rewritten in independent form to include the limitations of the base claim and any intervening claims. Applicant respectfully traverses these rejections and requests that these rejections be withdrawn for at least the reasons presented below.

Modha discusses systems and methods for clustering hypertext documents that a database maintains. The process of Modha "clusters hypertext document using words (contained in the document), out-links (from the document) and in-links (to the document)." Col 2, ln. 65 - Col. 3, ln. 1. Clustering in accordance with Modha provides a summary of hypertext documents returned from a search. Col. 1, lns. 28-30. The method begins with a search of hypertext documents in a database in response to a query to locate query result documents. A word dictionary is constructed using words within the query result documents, function words are pruned from the dictionary and a vector is formed for words remaining in the dictionary. The process is repeated for the out-links and in-links for a given query result document. Col. 1, lns. 33-58; Col. 4, ln. 6 - Col. 6, ln 45.

The vectors are normalized to create vector triplets that identify documents remaining in the dictionaries, which are clustered according to a toric k-means process. Col. 2, lns. 58-67. Modha discusses that the toric k-means process comprises an arbitrary segregation of the triplets into clusters, computing a set of triplets that describe the cluster and re-segregating by

putting each triplet into the cluster that it is most similar to. Col. 6, 48 - Col. 10, ln. 44. The process is repeated “until coherence of the obtained clusters no longer significantly increases,” and the clusters are annotated “using nuggets of information.” Col. 2, lns. 1-5.

By contrast to Modha, independent claim 1 of the present invention describes a computer-implemented method for generating superunits from a concept network, the concept network including a plurality of units (which may represent query terms) and a plurality of relationships defined between pairs of the plurality of units, each relationship having an associated edge weight. The method according to independent claim 1 comprises identifying a superunit seed comprising at least one member unit, wherein each member unit is one of the plurality of units in the concept network, and defining a signature for the superunit seed, the signature including one or more signature units, wherein each signature unit has a relationship has a relationship in the concept network with at least a minimum number of the member units. The method continues with expansion of the superunit seed by adding one or more new member units from the concept network, wherein each new member unit satisfies a match criterion based on the signature, and modifying the signature based on the expanded superunit seed. The acts of expanding and modifying are repeated until a convergence criterion is satisfied, wherein a final superunit and a final signature are formed and superunit membership information is stored for each member unit of the final superunit. Claims 29 and 40 comprise elements that are substantially similar to independent claim 1, bust cast as a system and computer program product, respectively.

Contrary to the Examiner’s assertion, Modha does not teach or suggest a concept network, which includes a plurality of units and a plurality of relationships between pairs of units, with a given relationship having an associated edge weight. The examiner indicates that the relationship between units reads onto the link dictionary of Modha. The link dictionary, however, identifies pages in the database that link to documents in a result set or to which

documents in the result set link. There is no teaching or suggestion that the link dictionary of Modha maintains a weight for a relationship between units. Furthermore, a concept network, as the term indicates, is a network of concepts (which may be represented by query terms) and not a database of hypertext documents, a given hypertext document providing no information with regard to the concepts contained in a given web page. Indeed, although Modha discusses word analysis of the hypertext documents, there is no teaching or suggestion to create a network of concepts comprising a plurality of units and a plurality of relationships between units; a given relationship having an associated edge weight.

Modha also fails to teach or suggest defining a signature for the superunit seed, the signature including one or more signature units each having a relationship in the concept network with at least a minimum number of member units. Modha discusses the geometric representation of documents in a result set by a triplet of unit vectors (D, F, B) for a given document. Col. 4, lns. 20-29. The triplet captures “information represented by the words contained in the document, the out-links originating at the document, and the in-links terminating at the document, respectively.” Id. A given triplet, however, does not comprise one or more signature units, wherein each signature unit has a relationship in the concept network with at least a minimum number of member units. Similarly, the claimed signature is defined for a superunit seed (which comprises at least one member unit from the units of the concept network) in relation to the concept network, and is unrelated to finding relationships between hypertext documents as Modha discusses.

Modha also fails to teach or suggest expanding the superunit seed as claimed, as well as modifying the signature based on expansion of the superunit seed. Indeed, the Examiner fails to put forward a position with regard to the step of modifying the signature based on the expanded superunit seed and Applicants respectfully assert that Modha is silent with regard to this point. Finally, Modha does not teach or suggest storing superunit membership for each

member unit of the final superunit. The portions of Modha upon which the Examiner relies only discuss annotating clusters “using nuggets of information, the nuggets including summary, breakthrough, review, keyword, citation and reference.” Col. 2, lns. 3-5. Annotating a cluster with “nuggets of information,” however, fails to teach or suggest storing superunit membership information for each member unit of the final superunit, as claimed.

At most, Modha discusses systems and techniques for clustering by determining the similarity between a plurality of hypertext documents. Col. 3, lns. 24-30. There are a number of clustering techniques known to those of skill in the art and presently before the Examiner. None of these techniques, however, including those that Modha discusses, teach or suggest the method of independent claim 1. Accordingly, Applicants respectfully request the withdrawal of the rejection of independent claim 1 as anticipated by Modha and allowance regarding the same. Applicants further assert that independent claims 29 and 40 are allowable over Modha for at least the reasons presented in conjunction with independent claim 1 and respectfully request allowance regarding the same.

The dependent claims of the present application contain additional features that further substantially distinguish the invention of the present application over Modha and the prior art of record. Given the applicants’ position on the patentability of the independent claims, however, it is not deemed necessary at this point to delineate such distinctions.

For at least all of the above reasons, Applicants respectfully request that the Examiner withdraw all rejections and objections, and allowance of all the pending claims is respectfully solicited. To expedite prosecution of this application to allowance, the examiner is invited to call the Applicants' undersigned representative to discuss any issues relating to this application.

Dated: April 5, 2007

THIS CORRESPONDENCE IS BEING
SUBMITTED ELECTRONICALLY THROUGH
THE PATENT AND TRADEMARK OFFICE EFS
FILING SYSTEM ON APRIL 5, 2007.

Respectfully submitted,



Seth H. Ostrow
Reg. No. 37,410
DREIER LLP
499 Park Ave.
New York, New York 10022
Tel : (212) 328-6100
Fax: (212) 328-6101

Customer No. 61834